

PENGEMBANGAN TES DIAGNOSTIK *TWO-TIER* DAN MANFAATNYA DALAM MENGUKUR KONSEPSI KIMIA SISWA SMA

Nahadi, Wiwi Siswaningsih & Rose Purnamasari

Jurusan Pendidikan Kimia, FPMIPA, Universitas Pendidikan Indonesia

Abstract: Development of Diagnostic Test TWO-TIER and Its Advantage to Measure Chemistry Conception of Senior High School Student. This study aims to develop a two-tier multiple choice diagnostic test as a tool to measure students' understanding of concepts in solubility and solubility product subject. The method used in this research is Research and Development. The research was conducted in two phases, the preliminary studies and model development. The test was developed by 30 items and tested to 28 high school students in XI grade. The test has high reliability and sufficient validity. 12 of 30 items were developed have good criteria and the rest still need improvement. Based on the questionnaire, students responded quite well to the developed test and most of the students like this kind of two-tier multiple choice diagnostic test compared to the usual tests.

Abstrak: Pengembangan Tes Diagnostik *Two-Tier* dan Manfaatnya dalam Mengukur Konsepsi Kimia Siswa SMA. Penelitian ini bertujuan untuk mengembangkan tes *diagnostik two-tier* sebagai alat untuk mengetahui penguasaan konsep kelarutan dan hasil kali kelarutan siswa SMA. Metode yang digunakan dalam penelitian ini adalah metode *Research and Development*. Penelitian ini dilakukan melalui dua tahap yaitu studi pendahuluan dan pengembangan model. Tes yang dikembangkan adalah sebanyak 30 butir soal dan diujicobakan kepada 28 siswa SMA kelas XI. Tes yang dikembangkan memiliki reliabilitas dengan kategori tinggi dan validitas tes dengan kategori cukup. Dari 30 butir soal, butir soal yang memenuhi kriteria sebagai alat ukur yang baik adalah sebanyak 12 butir soal dan selebihnya masih memerlukan perbaikan. Berdasarkan data angket, siswa memberikan respon cukup baik terhadap tes yang dikembangkan dan sebagian besar siswa menyukai jenis tes *diagnostik two-tier multiple choice* dibandingkan dengan tes biasa.

Keywords: Two-Tier Multiple Choice Diagnostic Test, Students' Understanding of Concept, Solubility and Solubility Product

PENDAHULUAN

Pembelajaran bertujuan untuk memberikan kemampuannya yang utuh kepada peserta didik. Penguasaan konsep merupakan hal yang penting dalam pembelajaran. Salah satu tujuan dari pembelajaran kimia adalah memahami konsep, prinsip, hukum, dan teori kimia serta saling keterkaitannya dan penerapannya untuk menyelesaikan masalah dalam kehidupan sehari-hari dan teknologi. Berdasarkan tujuan pembelajaran tersebut, pembelajaran kimia menuntut siswa untuk dapat memahami konsep-konsep kimia. Pada kenyataannya tidak semua siswa yang mengikuti proses pembelajaran dapat memahami konsep kimia yang dipelajarinya. Siswa sering mengalami kesulitan dalam memahami konsep-konsep kimia (Salirawati, 2010).

Beberapa cara yang digunakan untuk mengetahui pemahaman konsep siswa, diantaranya adalah dengan penggunaan peta konsep, wawancara dan tes diagnostik *two-tier*

multiple choice (Tuysuz, 2009). Menurut Tamir (Treagust, 2006; Chandrasegaran *et al.*, 2007), tes diagnostik *two-tier multiple choice* merupakan alat diagnostik yang efektif. Tes diagnostik *two-tier multiple choice* merupakan salah satu tes diagnostik yang mana soalnya merupakan soal bertingkat dua. Tingkat pertama terdiri dari pertanyaan dengan lima pilihan jawaban, sedangkan tingkat kedua terdiri dari lima pilihan alasan yang mengacu pada jawaban pada tingkat pertama. Alasan tersebut terdiri dari satu jawaban benar dan distraktor. Jawaban distraktor merupakan penjelasan siswa yang diperoleh dari literatur, *interview* ataupun dari respon terbuka (Tuysuz, 2009).

Two-tier multiple choice memiliki kelebihan dibandingkan dengan bentuk soal lain. Kelebihan *two-tier multiple choice* dibandingkan dengan *multiple choice* konvensional salah satunya adalah mengurangi *error* dalam pengukuran, dengan menggunakan

multiple choice konvensional dengan lima pilihan jawaban memiliki kesempatan menjawab benar dengan cara menebak adalah 20% sedangkan jika menggunakan tes *two-tier multiple choice* kesempatan menjawab benar dengan cara menebak adalah 4% (Tuysuz, 2009). Selain itu, dengan menggunakan tes diagnostik *two-tier multiple choice* guru akan lebih mudah dalam melakukan penskoran (Tuysuz, 2009). Hal serupa diungkapkan oleh Tan dan Treagust (1999), yang menyatakan bahwa tes diagnostik *two-tier multiple choice* lebih mudah dilaksanakan dan diberi skor dibandingkan dengan alat diagnostik lain, sehingga memberikan manfaat lebih bagi guru di kelas.

Tes diagnostik *two-tier multiple choice* telah dikembangkan dan digunakan untuk mengetahui pemahaman konsep siswa pada beberapa ilmu sains seperti biologi, kimia dan fisika. Penelitian mengenai tes diagnostik *two-tier multiple choice* telah dikembangkan pada beberapa pokok bahasan kimia, seperti pokok bahasan ikatan kovalen dan struktur oleh Peterson dan Treagust (1988), analisis kualitatif oleh Tan, dkk (2002), reaksi kimia oleh Chandrasegaran, Tregust dan Mocerino (2005), dan energi ionisasi unsur oleh Tan, dkk (2005). Di Indonesia, penelitian mengenai pengembangan tes diagnostik *two-tier multiple choice* dikembangkan oleh Ariyanti pada materi senyawa karbon (2012).

Data hasil wawancara dengan beberapa guru SMA di beberapa sekolah di kota Bandung, menunjukkan bahwa pada umumnya guru tidak menggunakan tes diagnostik untuk mengukur pemahaman konsep siswa. Pada umumnya guru memperoleh informasi mengenai pemahaman konsep siswa dari hasil ulangan harian dan alat ukur yang digunakannya berupa bentuk soal pilihan berganda biasa atau *essay*. Mengingat pentingnya pemahaman konsep siswa dalam pembelajaran, maka sebaiknya perlu dikembangkan tes diagnostik di sekolah.

Materi yang dipilih pada penelitian ini adalah materi kelarutan dan hasil kali kelarutan. Kebanyakan siswa menganggap bahwa materi kelarutan dan hasil kali kelarutan merupakan materi yang sulit (Onder, 2006). Dan menurut Sudijono (2007), materi yang ditanyakan dalam tes diagnostik pada umumnya ditekankan pada materi yang biasanya sulit menurut pengalaman siswa. Dengan demikian, tes diagnostik ini dikembangkan pada materi kelarutan dan hasil

kali kelarutan.

METODE PENELITIAN

Metode penelitian yang digunakan dalam penelitian ini adalah metode R&D (*Research and Development*) atau penelitian dan pengembangan. Menurut Sukmadinata (2008), metode R&D merupakan suatu proses atau langkah-langkah untuk mengembangkan suatu produk baru atau menyempurnakan produk yang telah ada yang dapat dipertanggungjawabkan. Secara garis besar langkah penelitian dan pengembangan hasil modifikasi yang dikembangkan oleh Sukmadinata dkk terdiri dari tiga tahap, yaitu: studi pendahuluan, pengembangan model dan uji model. Penelitian ini hanya dilakukan sampai tahap pengembangan model yaitu sampai langkah uji coba terbatas.

Objek penelitian adalah butir soal *tes diagnostik two-tier multiple choice* sebanyak 30 butir soal yang dikerjakan oleh 28 orang siswa SMA kelas XI di salah satu SMA di Kota Bandung yang telah memperoleh materi kelarutan dan hasil kali kelarutan. Tes diagnostik *two-tier multiple choice* ini dibuat melalui tiga tahap pembuatan soal, yaitu:

Soal Essay dan Wawancara

Tahap pertama adalah pembuatan soal *essay* dan dilengkapi oleh wawancara. Soal *essay* dibuat sebanyak 20 butir soal. Setelah dinilai valid secara isi, tes diujikan kepada satu kelas yang berjumlah 42 orang siswa. Dilengkapi pula dengan wawancara terhadap 12 orang siswa yang masing-masing terdiri dari empat orang kelompok rendah, empat orang kelompok sedang dan empat orang kelompok tinggi. Jawaban-jawaban yang diperoleh dari soal *essay* ataupun wawancara digunakan sebagai pilihan jawaban (*first tier*) pada soal tahap dua (soal pilihan ganda beralasan).

Soal Pilihan Ganda Beralasan

Tahap kedua adalah pembuatan soal pilihan ganda beralasan. Soal dibuat sebanyak 20 butir soal. Setelah dinilai valid secara isi, tes diujikan kepada kelas yang sama yang berjumlah 42 orang siswa. Soal ini terdiri dari lima pilihan jawaban yang terdiri dari satu jawaban benar dan empat jawaban pengecoh yang diperoleh baik dari jawaban soal *essay*,

wawancara ataupun dari literatur. Pada soal ini siswa dituntut untuk menjawab satu pilihan jawaban yang paling tepat dan menjelaskan alasan memilih jawaban tersebut. Alasan-alasan yang diungkapkan oleh siswa kemudian dijadikan sebagai pilihan alasan (*second tier*) pada soal tes diagnostik *two-tier multiple choice*.

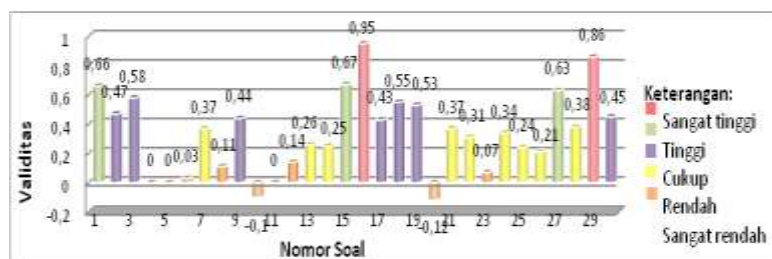
Soal Tes Diagnostik Two-Tier Multiple Choice

Tahap ketiga merupakan tahap pembuatan soal tes diagnostik *two-tier multiple choice*. Soal tes diagnostik *two-tier multiple choice* merupakan soal bertingkat dua. Tingkat pertama terdiri dari pertanyaan dengan lima pilihan jawaban yang diperoleh dari jawaban *essay*, wawancara ataupun literatur. Sedangkan tingkat kedua terdiri dari lima pilihan alasan yang mengacu pada jawaban pada tingkat pertama.

Jumlah soal yang dikembangkan adalah sebanyak 30 butir soal. Setelah dinilai valid secara isi, tes diujicobakan pada kelas yang sama yang berjumlah 28 siswa. Tes ini juga digunakan untuk mengukur pemahaman konsep siswa pada materi kelarutan dan hasil kali kelarutan.

Alur penelitian yang dilakukan adalah sebagai berikut:

1. Studi kepustakaan
2. Survei lapangan
3. Pembuatan kisi-kisi soal
4. Penyusunan produk awal
5. Uji validasi isi
6. Revisi butir soal
7. Uji coba terbatas
8. Pemberian angket
9. Pengolahan dan analisis data



Gambar 1. Gambar Validitas tiap Butir Soal

Dari Gambar 1 diketahui bahwa butir soal dengan kategori validitas sangat tinggi

Pengolahan data hasil uji coba tes diagnostik *two-tier multiple choice* meliputi perhitungan nilai validitas, reliabilitas, taraf kesukaran, daya pembeda, dan kualitas pengecoh. Pemahaman konsep siswa pada materi kelarutan dan hasil kali kelarutan diketahui dengan menafsirkan persentase skor siswa dan menganalisis pola jawaban siswa pada tes diagnostik *two-tier multiple choice*. Data angket digunakan untuk mengetahui bagaimana tanggapan siswa terhadap tes yang dikembangkan dan digunakan untuk memperkuat data hasil penelitian.

HASIL DAN PEMBAHASAN

Tes diagnostik *two-tier multiple choice* yang telah dinyatakan valid saat uji validitas isi kemudian diujicobakan. Dari hasil uji coba terbatas diperoleh data pola jawaban dan skor siswa pada tes diagnostik *two-tier multiple choice*. Data tersebut digunakan untuk menghitung validitas, reliabilitas, daya pembeda, taraf kesukaran dan indeks pengecoh yang kemudian digunakan untuk menentukan kualitas soal yang dikembangkan. Selain itu, data skor dan pola jawaban siswa digunakan untuk mengetahui pemahaman konsep siswa pada materi kelarutan dan hasil kali kelarutan. Berikut adalah data yang diperoleh dari penelitian ini:

Validitas

Pada penelitian ini validitas butir soal dihitung dengan menggunakan rumus korelasi *product-moment*. Validitas empiris yang digunakan adalah validitas *concurrent*. Hasil penafsiran validitas tiap butir soal dapat dilihat pada Gambar 1 sedangkan persentase kategori validitas tes secara keseluruhan dapat dilihat pada Gambar 2 di bawah ini.



Gambar 2. Gambar Persentase Kategori Validitas Butir Soal

sebesar 7%, tinggi 10%, cukup 23%, rendah 30% dan sangat 30%. Nilai validitas tes secara

keseluruhan adalah sebesar 0.41. Berdasarkan pedoman penafsiran yang dikemukakan Arifin (2009), nilai validitas 0.41 termasuk ke dalam kategori cukup. Faktor seperti siswa saling mencontek, melakukan penebakan saat menjawab, soal baru satu kali diujicobakan dan masih memerlukan revisi kemungkinan menjadi penyebab nilai validitas kecil pada penelitian ini.

Reliabilitas

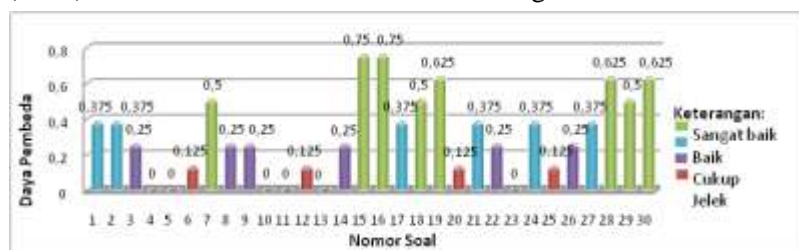
Nilai reliabilitas keseluruhan untuk tes yang dikembangkan adalah sebesar 0.70. Berdasarkan pedoman penafsiran reliabilitas yang dikemukakan oleh Arifin (2009), tes tersebut termasuk ke dalam kategori

reliabilitas tinggi.

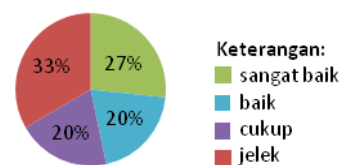
Daya Pembeda

Daya pembeda dari tes yang dikembangkan memiliki kategori yang bervariasi yaitu sangat baik, baik, cukup dan jelek. Nilai daya pembeda tiap butir soal beserta peafsirannya dapat dilihat pada Gambar 3 sedangkan persentase kategori daya pembeda pada tes secara keseluruhan dapat dilihat pada Gambar 4 di bawah ini.

Dari Gambar 4 dapat terlihat bahwa butir soal dengan kategori sangat baik adalah sebesar 27%, kategori baik 20%, kategori cukup 20%, dan kategori jelek 33%.



Gambar 3. Diagram Daya Pembeda tiap Butir Soal



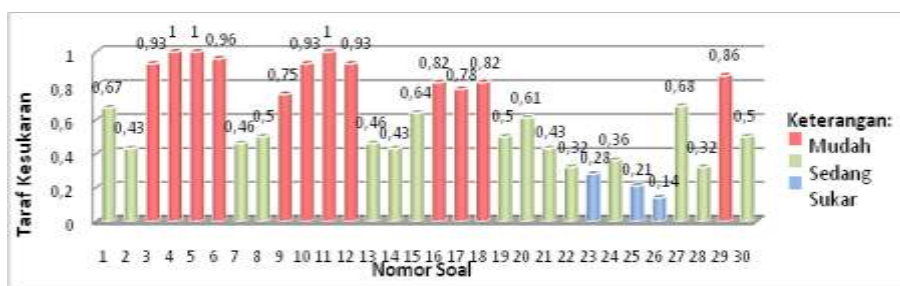
Gambar 4. Diagram Persentase Kategori Daya Pembeda Butir Soal

Taraf Kesukaran

Soal yang baik adalah soal yang memiliki tingkat kesukaran yang seimbang (Arifin, 2009). Oleh karena itu, dalam penyusunan soal perlu memperhatikan taraf kesukarannya. Data taraf kesukaran tiap butir soal yang beserta penafsirannya dapat dilihat pada Gambar 5 sedangkan persentase kategori taraf kesukaran pada tes secara keseluruhan

dapat dilihat pada Gambar 6 di bawah ini.

Butir soal dengan kategori mudah adalah sebesar 40%, kategori sedang sebesar 50%, dan kategori sukar sebesar 10%. Taraf kesukaran soal pada tes yang dikembangkan belum memiliki proporsi soal yang seimbang. Agar proporsi soal seimbang sebagian soal dengan kategori mudah harus ditingkatkan kesulitan menjadi sukar.



Gambar 5. Diagram Taraf Kesukaran Tiap Butir Soal

Tabel 1. Kategori Kualitas Pengecoh First Tier dan Second Tier

No Soal	Pilihan Jawaban First Tier					No Soal	Pilihan Jawaban Second Tier				
	A	B	C	D	E		A	B	C	D	E
1	SJ	SB	**	J	J	1	B	**	J	KB	J
2	SJ	**	KB	J	J	2	**	SJ	J	J	SB
3	J	SJ	J	**	J	3	**	SJ	SB	SB	J
4	J	J	**	J	J	4	J	SJ	J	**	J

5	J	J	J	J	**
6	J	SJ	J	J	**
7	B	**	B	B	J
8	J	KB	SJ	**	J
9	J	SB	SJ	J	**
10	J	J	**	J	J
11	SJ	**	J	J	J
12	J	J	J	J	**
13	J	J	SJ	J	**
14	**	J	J	SJ	SB
15	KB	SJ	J	**	J
16	**	J	SJ	J	SB
17	SJ	**	J	KB	J
18	SB	**	SJ	J	J
19	KB	SJ	SB	**	J
20	J	B	**	B	SJ
21	SJ	J	**	B	B
22	SJ	J	J	**	B
23	**	B	KB	KB	KB
24	J	B	B	**	J
25	SB	B	KB	J	**
26	**	SJ	J	SB	J
27	B	B	**	B	B
28	KB	**	J	B	SJ
29	**	J	SB	J	SB
30	**	J	SJ	J	J

5	J	J	J	**	J
6	J	SJ	**	J	J
7	SJ	J	J	KB	**
8	J	SJ	J	**	J
9	SJ	J	**	J	J
10	J	**	SJ	J	J
11	**	J	J	J	J
12	J	J	J	J	**
13	**	SJ	KB	B	J
14	SJ	**	B	J	J
15	J	KB	**	KB	SJ
16	SJ	J	**	J	SB
17	**	J	J	SB	SJ
18	**	J	SB	J	SJ
19	J	SJ	**	SB	KB
20	J	J	J	J	**
21	**	J	KB	SJ	J
22	**	SJ	J	B	J
23	KB	**	J	J	J
24	SJ	KB	KB	**	J
25	J	KB	J	KB	**
26	J	J	**	J	J
27	KB	**	B	SB	B
28	B	J	SJ	**	B
29	J	**	J	B	B
30	J	**	B	SJ	J

Indeks Pengecoh

Kualitas soal pilihan ganda selain ditentukan oleh validitas, reliabilitas, daya pembeda dan taraf kesukaran juga ditentukan oleh kualitas pengecoh. Butir soal yang baik adalah butir soal yang pengecohnya dipilih secara merata oleh siswa yang menjawab salah (Arifin, 2009).

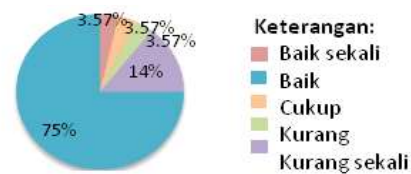
Kualitas pengecoh pada tes yang dikembangkan dibagi menjadi dua macam yaitu kualitas pengecoh pada soal tingkat pertama (*first tier*) dan pada tingkat kedua (*second tier*). Data kualitas pengecoh tiap pilihan jawaban pada tiap butir soal beserta penafsirannya dapat dilihat pada Tabel 1.

Berdasarkan data pada Tabel 1 dapat dilihat bahwa kualitas pengecoh pada tes yang dikembangkan meliputi kategori sangat baik, baik, kurang baik, jelek dan sangat jelek. Sebanyak 20 butir soal masing-masing pada *first* dan *second tier* memiliki satu pengecoh dengan kategori sangat jelek dan sebagian besar

pengecoh pada butir soal yang dikembangkan masuk dalam kategori jelek.

Tingkat Pemahaman Konsep Siswa

Persentase kategori tingkat pemahaman siswa pada tes secara keseluruhan dapat dilihat pada Gambar 7.



Gambar 7. Diagram Persentase Tingkat Pemahaman Siswa

Sebagian besar siswa memiliki tingkat pemahaman yang baik pada materi kelarutan dan hasil kali kelarutan. Namun, bukan berarti bahwa siswa memahami seluruh konsep materi tersebut.

Berdasarkan analisis terhadap pola

jawaban siswa pada tiap butir soal, dapat diketahui bahwa ada siswa yang telah memahami, tidak memahami bahkan ada siswa yang mengalami miskonsepsi. Kategori tingkat pemahaman siswa pada materi kelarutan dan hasil kali kelarutan berdasarkan pola jawabannya pada tes diagnostik two-tier multiple choicedapat dilihat pada Tabel 2.

Dari data persentase kategori tingkat pemahaman siswa untuk setiap butir soal dapat diketahui materi kelarutan dan hasil kali kelarutan mana yang sudah dipahami siswa, tidak dipahami, dipahami sebagian, dan miskonsepsi. Sehingga dari kategori tersebut dapat disimpulkan pada materi mana saja siswa mengalami kesulitan. Pada umumnya siswa mengalami kesulitan pada teori/konsep kelarutan dan hasil kali kelarutan dibandingkan dengan soal hitungan. Hal ini kemungkinan

disebabkan oleh proses pembelajaran yang diterima oleh siswa, yang lebih menekankan pada aplikasi hitungan daripada konsep/teori. Berdasarkan hasil analisis pola jawaban siswa, sebagian besar siswa mengalami kesulitan pada konsep hubungan antara pH dengan kelarutan. Adapun kesulitan tersebut dapat disebabkan oleh miskonsepsi yang terjadi ataupun karena siswa tidak memahami konsep tersebut.

Data Angket

Data respon siswa pada tiap pertanyaan dalam angket dapat dilihat pada Tabel 3. Pertanyaan angket dapat diklasifikasikan menjadi beberapa kategori pertanyaan, yaitu: Pertanyaan nomor 1 dan 2 menanyakan keterbacaan soal, apakah pertanyaan dan alasan

Tabel 2. Kategori Tingkat Pemahaman Siswa pada Materi Kelarutan dan Hasil Kelarutan

No Butir Soal	Kategori Tingkat Pemahaman(%)						
	M	Mi-1	Mi-2	TM-1	TM-2	MS-1	TM-3
1	71.4	3.6	0	25	0	0	0
2	46.4	25	10.7	14.3	0	3.6	0
3	92.9	7.1	0	0	0	0	0
4	100	0	0	0	0	0	0
5	100	0	0	0	0	0	0
6	96.4	0	0	3.6	0	0	0
7	46.4	42.9	0	10.7	0	0	0
8	50.0	28.6	0	21.4	0	0	0
9	71.4	7.1	14.3	7.1	0	0	0
10	92.8	0	0	3.6	0	0	3.6
11	100	0	0	0	0	0	0
12	96.4	0	3.6	0	0	0	0
13	46.4	14.3	0	39.3	0	0	0
14	42.9	0	10.7	46.4	0	0	0
15	60.7	21.4	3.6	10.7	0	3.6	0
16	82.1	0	3.6	14.3	0	0	0
17	78.6	3.6	3.6	14.3	0	0	0
18	78.6	3.6	3.6	14.3	0	0	0
19	57.1	21.4	3.6	17.9	0	0	0
20	64.3	0	21.4	14.3	0	0	0
21	46.4	0	0	53.6	0	0	0
22	35.7	3.6	21.4	39.3	0	0	0
23	28.6	28.6	0	42.9	0	0	0
24	32.1	3.6	50.0	14.3	0	0	0
25	28.6	0	7.1	64.3	0	0	0
26	14.3	25.0	39.3	21.4	0	0	0
27	75.0	7.1	7.1	10.7	0	0	0
28	28.6	7.1	3.6	60.7	0	0	0
29	89.3	0	3.6	7.1	0	0	0
30	53.6	10.7	32.1	3.6	0	0	0

Tabel 3. Respon Siswa terhadap Tes Diagnostik Two-Tier Multiple Choice

No. Pertanyaan	Persentase Jawaban Siswa (%)	
	Ya	Tidak
1	42.9	57.1
2	82.1	17.9
3	100	0
4	61.5	38.5
5	85.7	14.3
6	92.9	7.1
7	46.2	53.8
8	14.3	85.7
9	37	63
10	92.9	7.1
11	85.7	14.3
12	75	25
13	72	28

jawaban dapat dipahami. Responden yang menyatakan bahwa pertanyaan dapat dipahami adalah sebesar 42.9% dan responden yang menyatakan bahwa alasan jawaban sulit dipahami adalah sebesar 82.1%. Pertanyaan nomor 3 menanyakan jawaban (first tier) dan alasan jawaban (second tier) saling berkaitan. Semua responden setuju bahwa jawaban dan alasan jawaban saling berkaitan. Pertanyaan nomor 4 menanyakan tingkat kesulitan materi kelarutan dan hasil kali kelarutan. Sebanyak 61.5% responden menyatakan bahwa materi kelarutan dan hasil kali kelarutan adalah materi yang sulit. Pertanyaan nomor 5 dan 6 menanyakan apakah ada soal yang dianggap mudah dan sulit pada tes yang dikembangkan. Sebanyak 85.7% responden setuju bahwa ada soal yang dianggap mudah dan sebanyak 92.9% responden setuju ada soal yang dianggap sulit. Pertanyaan nomor 7 menanyakan apakah soal yang biasa diberikan untuk materi kelarutan dan hasil kali kelarutan hanya berupa soal hitungan. Sebanyak 53.8% responden menjawab tidak. Pertanyaan nomor 8 menanyakan apakah siswa menyukai soal mengenai teori daripada soal hitungan. Sebanyak 85.7% responden menyatakan lebih menyukai soal berupa hitungan. Pertanyaan nomor 9 menanyakan apakah siswa sebelumnya pernah mengerjakan soal tes diagnostik two-tier multiple choice. Sebanyak 63% responden menyatakan belum pernah mengerjakan soal tes diagnostik two-tier multiple choice. Pertanyaan nomor 10 menanyakan apakah soal tes diagnostik two-tier multiple choice dapat menggali pemahaman siswa terhadap materi yang diujicobakan.

Responden sebanyak 92.9% menyatakan setuju. Pertanyaan nomor 11

menanyakan apakah terdapat kendala dalam mengerjakan soal tes diagnostik two-tier multiple choice. Responden sebanyak 85.7% menyatakan mengalami kendala. Pertanyaan nomor 12 menanyakan apakah waktu yang diberikan untuk mengerjakan tes cukup. Responden sebanyak 75% menyatakan waktu yang diberikan untuk mengerjakan soal adalah cukup. Pertanyaan nomor 13 menanyakan respon siswa terhadap penggunaan soal tes diagnostik two-tier multiple choice. Sebanyak 72% responden lebih menyukai soal two-tier multiple choicedibandingkan soal yang biasa diberikan.

Berdasarkan analisis terhadap hasil penelitian, diketahui bahwa dari 30 butir soal yang dikembangkan, sebanyak 12 butir soal dianggap telah memenuhi kategori sebagai soal yang baik dilihat dari validitas, reliabilitas dan daya pembedanya. Namun, dilihat dari indeks pengecohnya butir soal tersebut masih memerlukan perbaikan.

Pemahaman konsep siswa terhadap materi kelarutan dan hasil kali kelarutan sebagian besar sudah tergolong baik. Hal ini menunjukkan bahwa sebagian besar siswa mampu menjawab sebagian besar pertanyaan yang disajikan dalam tes. Ini dapat dilihat dari taraf kesukaran soal yang menunjukkan bahwa sebanyak 40% butir soal termasuk ke dalam kategori mudah. Namun demikian, masih terdapat beberapa siswa yang mengalami kesulitan untuk memahami materi kelarutan dan hasil kali kelarutan. Kesulitan yang terjadi dapat disebabkan oleh miskonsepsi ataupun karena tidak memahami konsep yang ditanyakan.

Dalam pelaksanaan tes khususnya tes kelarutan dan hasil kali kelarutan, siswa lebih menyukai pertanyaan hitungan daripada teori. Hal tersebut dapat dilihat dari tingkat kesalahan siswa saat menjawab pertanyaan mengenai teori lebih besar daripada pertanyaan mengenai hitungan. Siswa cenderung mengalami kesulitan dalam menjawab pertanyaan mengenai teori/konsep khususnya teori mengenai hubungan antara pH dengan kelarutan. Hal ini sejalan dengan hasil analisis taraf kesukaran siswa, butir soal yang menanyakan hubungan antara kelarutan dengan pH termasuk ke dalam kategori sukar.

SIMPULAN DAN SARAN

Berdasarkan hasil penelitian yang telah dilakukan diperoleh simpulan sebagai berikut:

1. Tes diagnostik *two-tier multiple choice* yang dikembangkan secara keseluruhan termasuk ke dalam kategori soal yang baik dilihat dari validitas, reliabilitas serta taraf kesukarannya. Tes yang dikembangkan memiliki reliabilitas dengan kategori tinggi (0.70) dan validitas tes dengan kategori cukup (0.41). Hasil analisis butir soal menunjukkan validitas dengan kategori sangat tinggi sebanyak 7%, kategori tinggi 10%, kategori cukup 23%, kategori rendah 30% dan kategori sangat rendah 30%. Daya pembeda butir soal yang termasuk ke dalam kategori sangat baik sebanyak 27%, kategori baik 20%, kategori cukup 20%, dan kategori jelek 33%. Tes ini memiliki taraf kesukaran yang bervariasi yaitu soal dengan kategori sukar sebanyak 10%, kategori sedang 50% dan kategori mudah 40%.
2. Berdasarkan analisis data, tes diagnostik *two-tier multiple choice* yang dikembangkan belum dapat mengukur pemahaman konsep siswa.
3. Respon siswa terhadap tes diagnostik *two-tier multiple choice* yang dikembangkan cukup baik dan siswa lebih menyukai jenis soal *two-tier multiple choice* dibandingkan dengan soal yang biasa diberikan.

DAFTAR RUJUKAN

Arifin, Z. 2009. *Evaluasi Pembelajaran*. Bandung: Rosda Karya.

Chandrasegaran, A. L, Treagust, D. F and Mocerino, M. 2007. "The Development of A Two-Tier Multiple-Choice Diagnostik Instrumen for Evaluating Secondary School Students' Ability to Describe and Explain Chemical Reactions Using Multiple Levels of Representation". *Chemistry Education Research and Practice*. 8, (3): 293--307.

Onder, I. 2006. *The Effect of Conceptual Change Approach on Students' Understanding of Solubility Equilibrium Concept*. Disertasi pada Middle East Technical University.

Salirawati, D. 2010. *Pengembangan Model Instrumen Pendeteksi Miskonsepsi Kimia pada Peserta Didik SMA*. Disertasi Doktor UNY. Bandung: tidak diterbitkan.

Sudijono, A. 2007. *Pengantar Evaluasi Pendidikan*. Jakarta: Raja Grafindo Persada.

Sukmadinata, N. S. 2008. *Metode Penelitian Pendidikan*. Bandung: Rosda Karya.

Tan, D. K. C. and Treagust, D. F. 1999. Evaluating Students' Understanding of Chemical Bonding. *School Science Review*, 81:75—.

Treagust, D. F. 2006. *Diagnostic Assessment in Science as a Means to Improving Teaching, Learning and Retention*. (Invited resentation). Uni Serve Science Assessment Symposium *Proceedings*.

Tuysuz, C. 2009. "Development of Two-Tier Diagnostic Instrument and Assess Students' Understanding in Chemistry". *Scientific Research and Essay*, 4(6): 626--631.